

BAB I

PENDAHULUAN

A. Latar Belakang Masalah

Perkembangan teknologi informasi dan komunikasi di Indonesia semakin meningkat seiring dengan perkembangan zaman. Perkembangan teknologi informasi dan komunikasi di Indonesia telah memberikan banyak manfaat pada bidang industri telekomunikasi. Industri telekomunikasi telah memasuki kondisi persaingan yang jauh lebih ketat dengan penyedia layanan komunikasi baik lokal maupun asing (Susetyo, 2010). Berbagai perusahaan penyedia jasa layanan komunikasi di Indonesia telah berkembang untuk memenuhi kebutuhan dan kepuasan masyarakat. PT Indonesian Satellite Corporation Tbk. (Indosat) merupakan salah satu perusahaan penyedia jasa layanan telekomunikasi di Indonesia. Indosat telah menjadi operator telepon selular yang terbesar kedua di Indonesia (Safarizky, 2010).

Indosat terus memperluas jaringan dan meningkatkan promosi pada produk yang ditawarkan untuk tetap dapat bersaing dengan penyedia layanan seluler yang lain. Tujuan utama promosi yang dilakukan yaitu untuk memberikan informasi dan menarik perhatian konsumen guna meningkatkan penjualan. Promosi merupakan faktor yang sangat penting untuk menentukan keberhasilan suatu program pemasaran. Produk yang dihasilkan oleh Indosat dapat dikenal oleh konsumen apabila melakukan promosi. Pelanggan akan mengetahui manfaat dan kualitas produk yang dihasilkan oleh Indosat melalui promosi. Promosi dapat

dilakukan melalui berbagai media komunikasi. Salah satu media komunikasi yang digunakan untuk mempromosikan produk layanan telekomunikasi yaitu Twitter.

Twitter merupakan salah satu media sosial yang berfungsi untuk mempermudah komunikasi masyarakat. Berdasarkan Pew Research Center (2018), Twitter menempati peringkat tujuh sebagai media sosial yang paling banyak digunakan oleh warga Amerika Serikat. Mayoritas pengguna Twitter berada pada kalangan remaja akhir, yaitu pada usia 18 tahun hingga 24 tahun. Dilansir dalam Semiocast (2012), Indonesia menempati posisi kelima dengan jumlah pengguna Twitter terbanyak dan Jakarta merupakan kota yang paling aktif melakukan postingan *tweet* di Twitter. Pengguna Twitter di Indonesia selalu meningkat pada setiap tahunnya. Dilansir dalam Statista (2016), Indonesia menempati posisi ketiga dengan jumlah pengguna Twitter aktif sebesar 24,34 juta pengguna pada bulan Mei tahun 2016.

Indosat telah memiliki sebuah akun Twitter resmi yang digunakan untuk mempromosikan produk layanan yang ditawarkan yaitu @IndosatCare. Cara mempromosikan produk layanan Indosat melalui Twitter yaitu dengan membuat *tweet* yang berisi tentang produk layanan Indosat yang ditawarkan. *Tweet* merupakan sebuah pesan yang berisi teks, foto, tautan atau video yang berkaitan dengan produk layanan yang ditawarkan. Pengguna dapat mengirim dan membaca *tweet* dengan batas maksimal 140 karakter. Pengguna juga dapat melihat *tweet* pengguna lain dan menemukan topik menarik yang sedang diperbincangkan. Adanya *tweet* yang berisi promosi pada akun Twitter resmi Indosat memungkinkan pengguna untuk menanggapi setiap postingan *tweet* dengan

memberikan masukan dan kritikan terhadap layanan Indosat. Opini yang diberikan masyarakat tidak hanya bersifat positif, melainkan juga bersifat negatif. Bahasa yang digunakan untuk memberikan opini dalam bentuk *tweet* tidak hanya menggunakan bahasa Indonesia, tetapi juga banyak yang menggunakan bahasa Inggris. Informasi yang terkandung di dalam *tweet* tersebut merupakan informasi penting karena dapat digunakan untuk mengetahui tingkat kepuasan masyarakat terhadap layanan Indosat dan faktor-faktor yang mempengaruhi tingkat kepuasan. Informasi tersebut berbentuk data opini yang dapat diolah dengan menggunakan *text mining*.

Text mining merupakan pencarian informasi yang dilakukan oleh komputer untuk mendapatkan informasi baru berdasarkan sumber data teks yang berbeda-beda (Rodiyansyah, 2012). Menurut Aji Suprpto (2017) salah satu tujuan *text mining* yaitu analisis sentimen. Analisis sentimen merupakan proses pengklasifikasian sebuah opini ke dalam opini positif maupun opini negatif. Salah satu jenis metode *machine learning* untuk analisis sentimen adalah *Naïve Bayes Classifier* (NBC). NBC merupakan metode sederhana yang berbasis probabilistik tetapi memiliki tingkat akurasi dan performansi yang tinggi dalam pengklasifikasian suatu teks (Routray, Swain, & Mishra, 2013).

Secara umum *text mining* harus melewati tahapan praproses dan analisis untuk memperoleh hasil yang diharapkan. Tahapan praproses terdiri atas *case folding*, *tokenizing*, penghapusan *stopwords*, normalisasi kata dan *stemming*. *Case folding* merupakan pengubahan semua karakter menjadi huruf kecil dalam dokumen. *Tokenizing* merupakan proses pemecahan dokumen teks menjadi token

atau kata. *Stopwords* merupakan daftar kata-kata yang dianggap tidak merepresentasikan konten dari suatu dokumen teks sehingga perlu dilakukan penghapusan. Normalisasi kata yaitu mengubah singkatan menjadi kata baku dan *stemming* merupakan pengubahan kata menjadi bentuk kata dasarnya. Setiap kata yang dihasilkan pada tahap praproses akan digunakan untuk proses analisis.

Vivek Narayanan, Ishan Arora & Arjun Bhatia (2013) melakukan penelitian analisis sentimen untuk mengetahui opini ulasan film oleh penonton India dengan menggunakan klasifikasi *Naïve Bayes*. Ulasan film pada penelitian ini bersumber dari *Internet Movie Database* (IMDb). Hasil penelitian cepat dan akurat dengan tingkat akurasi klasifikasi *Naïve Bayes* sebesar 88.80% dari data sebanyak 25.000 opini. Penelitian tentang analisis sentimen juga dilakukan oleh Rumana Rashid (2017). Analisis sentimen yang dilakukan yaitu untuk mengetahui opini konsumen terhadap layanan dua supermarket yaitu Walmart dan Target dengan menggunakan klasifikasi *Naïve Bayes* dan *Support Vector Machine*. Analisis sentimen penting dilakukan karena dapat digunakan sebagai salah satu penentu keputusan strategi pasar. Opini konsumen terhadap layanan kedua supermarket cenderung positif. Penelitian ini menghasilkan persentase dari prediksi opini positif yang benar dengan menggunakan klasifikasi *Naïve Bayes* yaitu sebesar 93.2% untuk supermarket Walmart dan 85% untuk supermarket Target sedangkan dengan menggunakan klasifikasi *Support Vector Machine* (SVM) yaitu sebesar 81.5% untuk supermarket Walmart dan 81% untuk supermarket Target.

Penelitian di Indonesia yang berkaitan dengan analisis sentimen dilakukan oleh Fahmi Nurulbaiti & Retno Subekti (2018). Penelitian ini dilakukan dengan

tujuan untuk mengetahui langkah-langkah dalam melakukan analisis sentimen masyarakat Indonesia dengan menggunakan Program R, mengetahui topik-topik yang sering diperbincangkan masyarakat Indonesia terkait dengan BPJS dan mengetahui tanggapan masyarakat terhadap keberadaan BPJS. Teknik analisis data pada penelitian ini dilakukan dengan *text mining* dan *scoring sentiment*. Hasil penelitian menunjukkan bahwa topik-topik yang paling sering dibicarakan masyarakat di jejaring sosial Twitter berkaitan dengan BPJS antara lain riba, dokter, tolak, kesehatan, dan pasien. Analisis sentimen *tweets* masyarakat pada jejaring sosial Twitter terhadap BPJS adalah sentimen negatif. Hal ini berarti program BPJS masih belum sepenuhnya memuaskan masyarakat.

Andita Wahyuningtyas (2016) juga melakukan penelitian tentang deteksi *tweet spam* dan bukan *spam* di jejaring sosial Twitter dengan mengklasifikasikan *spam* dan bukan *spam* menggunakan algoritma *Naïve Bayes*. Metode ini dipilih karena kesederhanaan algoritma yang mudah untuk diimplementasikan. Data Twitter yang diperoleh akan dilakukan tahap praproses yaitu *tokenizing*, *filtering*, normalisasi kata, *stemming*. Akurasi hasil klasifikasi *tweet spam* dan bukan *spam* adalah 95.57%. Berdasarkan hasil penelitian dapat diketahui kata-kata yang sering muncul pada kelas *spam* adalah bahasa, *follow*, inggris. Hal ini berarti *tweet spam* didominasi dengan konten *tweet* promosi bahasa inggris.

Mengingat tingkat kepentingan informasi yang terkandung dalam opini konsumen dan kelebihan dari metode *Naïve Bayes* untuk menggali informasi dari opini tersebut, penulis akan melakukan penelitian mengenai analisis sentimen dengan melakukan klasifikasi opini pengguna layanan Indosat di jejaring sosial

Twitter menggunakan algoritma *Naive Bayes Classifier*. Analisis sentimen dilakukan untuk mengetahui seberapa besar kepuasan pengguna layanan Indosat. Analisis sentimen data Twitter penting dilakukan karena hasil klasifikasi dapat dimanfaatkan sebagai evaluasi untuk meningkatkan kualitas pelayanan.

B. Batasan Masalah

Batasan masalah pada penelitian ini yaitu model klasifikasi hanya dilakukan untuk analisis opini berbahasa Indonesia.

C. Rumusan Masalah

Berdasarkan uraian latar belakang permasalahan di atas, maka diperoleh rumusan masalah sebagai berikut:

1. Bagaimana tahap-tahap klasifikasi dengan menggunakan algoritma *Naive Bayes Classifier* pada opini pengguna layanan Indosat di jejaring sosial Twitter?
2. Bagaimana hasil klasifikasi dari penerapan algoritma *Naive Bayes Classifier* pada opini pengguna layanan Indosat di jejaring sosial Twitter?

D. Tujuan

Tujuan penelitian ini adalah:

1. Mengetahui tahap-tahap klasifikasi dengan menggunakan algoritma *Naive Bayes Classifier* pada opini pengguna layanan Indosat di jejaring sosial Twitter.
2. Mengetahui hasil klasifikasi dari penerapan algoritma *Naive Bayes Classifier* pada opini pengguna layanan Indosat di jejaring sosial Twitter.

E. Manfaat

Manfaat penelitian ini antara lain adalah :

1. Bagi peneliti

Mengintegrasikan Algoritma *Naïve Bayes Classifier* sebagai metode untuk mendapatkan akurasi terbaik dalam klasifikasi opini.

2. Bagi pengguna

Pengguna dapat lebih mudah mengetahui keuntungan dan kerugian menggunakan layanan Indosat berdasarkan opini yang diberikan oleh pengguna Indosat.

3. Bagi Indosat

Indosat akan lebih mudah mengklasifikasikan opini pengguna layanan Indosat sehingga dapat dimanfaatkan sebagai alat bantu dalam pengambilan keputusan strategi pengembangan layanan dan produk. Hasil penelitian juga dapat dimanfaatkan sebagai evaluasi untuk meningkatkan kualitas pelayanan dan produk yang ditawarkan.